In the Specification

Please substitute the following paragraph on page 1, line 17:

Biocatalysis, defined as the biological synthesis of the molecules in question enzymatically, has been becoming more popular by offering a strong alternative to chemical synthesis, in terms of cost, time, purification steps, and simplicity of use. The introduction of any new biocatalysis process on an industrial scale necessitates, however, (i) identifying the enzyme (or the enzymes) which make(s) it possible to specifically convert the substrate provided into the desired product, (ii) identifying the enzyme (or the enzymes) which make(s) it possible to implement the catalysis in a stable manner and in the particular conditions linked to the industrial process (thermostability, pH, or tolerance to denaturation conditions of organic solvents~~, etc~~).

Please substitute the following paragraphs on page 2, line 31 to page 3, line 8:

However, the promising approach of exploiting these bacterial functions has always been considerably limited by a technological obstacle : the isolation and *in vitro* culture of the enormous potential offered by the bacterial diversity. Most bacteria developing in complex natural environments (soils and sediments, aquatic environments, digestive ~~systems, ...~~ systems) have not been cultivated because their optimal culturing conditions are unknown or too difficult to reproduce. Numerous scientific works demonstrate this established fact, and it is now widely admitted that only between 0.1 and 1% of the bacterial diversity, including all environments, have been isolated and cultivated (Amann *et al*, 1995, Microb. Rev., 59 :143-169). Even if the search for novel biocatalytic pathways within collections of microbic strains has proved to be effective, it nevertheless has the major disadvantage of only exploiting a tiny part of the bacterial biodiversity.

New approaches have been developed in order to overcome this critical point of isolating bacteria ~~and~~ in order to gain access to this enormous genetic potential offered by the adaptation systems of bacteria developed over their long evolution. This approach is called Metagenomics because it relates to a set of genomes from a bacterial community without any distinction (metagenome).

Please substitute the following paragraph on page 4, line 1:

The screening of enzymatic activities or of antibacterial activities from metagenomic libraries has been widely described in the scientific literature. The studies have related, for example, to the direct detection of chitinase (Cottrell *et al.*, 1999, Appl. Environ. Microbiol., 65 :2553-2557), lipase (Henne *et al.*, 2000, Appl. Environ. Microbiol.,66 :3113-3116), DNA, and amylase (Rondon *et al.*, 2000, Appl. Environ. Microbiol.,66 :2541-2547)~~ etc...~~ activity. In these studies, the host bacteria containing the recombinant clones are placed in culture on a medium complemented by the substrate of which metabolisation is sought, and the screening of the activity is generally based upon the appearance of haloes or precipitates around the colonies, or by a change to the appearance of the colonies which are metabolising the substrate being studied. It should be noted that the enzymatic activities detected by means of these examples are new activities for the host bacterium, but are not essential for the growth of the latter in the examples provided. A similar approach was described in the patent (Chromaxome~~ N°~~ No. 5,783,431). This patent describes a method of screening activity based upon the encapsulation of individual or pooled clones from a library in a stable, inert and porous matrix (advantageously alginate), in the form of macro- or micro-droplets. The droplets are for example subjected to a liquid culture containing the nutritive elements necessary for bacterial growth and a substrate (for example X-glucosaminide, X-acetate, X-glucopyranoside~~, etc..~~) the metabolisation of which is expressed by the appearance of blue colouring.

Please substitute the following paragraph on page 4, line 28:

All of these screening methods require the use of high throughput systems because they involve subjecting all of the clones to the screening test in order to identify the clones in question which respond positively to the tests. For this purpose, the company Diversa, leader in the domain of the discovery of new molecules, has developed a unique platform, called the GigaMatrix, enabling ultra-high throughput screening, of around 1 billion clones per day (~~http://www.diversa.com/techplat/gigamatrix/default.asp~~ See Worldwide Website: diversa.com/techplat/gigamatrix/default.asp).

Please substitute the following paragraph on page 5, line 1:

Another approach has already been described in patent WO 00/22170 of Microgenomics (~~№~~ US Patent No. 6,368,793 B1). This patent describes a methodology for identifying a metabolic pathway transforming a substrate S into a desired product T by creating or identifying a genetically manipulated organism of which the capability of implementing this reaction is placed under the control of an inducible promoter. This organism is used for screening fragments of nucleic acids in order to detect a gene involved in the transformation of a substrate into a product. The implementation of this method requires the identification and genetic characterisation of the genes responsible for the degradation of T in the expression host so that they can be placed under the control of an inducible promoter. This type of construct ~~can not~~ cannot always be considered, in particular when the genes in question are spread over the genome and there is a possible risk of "leaking" into the inducer. On the other hand, it represents extremely hard work which has to be repeated for every study of a product T. Finally, in this approach, the organism used must be capable of incorporating and metabolising S and T. All of the elements mentioned demonstrate the limits of the efficacy of this type of approach.

Please substitute the following paragraph on page 5, line 33 through page 6, line 2:

The time required and the means used to create the metagenomic library and then its screening ~~are~~ is therefore key, with small hope of success. An increase in the chances of discovery would involve, absolutely, the creation of a metagenomic library specific to each problem, in order to best respond to the objectives set.

Please substitute the following paragraph on page 13, line 24:

- transformed clones resulting from mutation, of phenotype (Ai- ; B-), having lost the capability of growing on {Ai} and on {B}. This change of phenotype can be explained either because (i) the metabolic pathway of {Ai} passes via {B} and the metabolism of {B} is disrupted (mutated phenotype IIIa), or because (ii) the mutagenesis has reached an element common to {Ai} and {B} such as for example a regulation element, a common transporter ~~etc.~~ (phenotype IIIb).

Please substitute the following paragraph on page 15, line 17:

The invention further relates to the selection of a host cell incapable of metabolising a substrate {Ai} and incapable of metabolising a desired product {B}. Preferably, this host cell is a bacterial host. These are, non-restrictively, *E. coli*, *Bacillus*, *Streptomyces*, *Pseudomonas*, and *Nocardia*, *Acinetobacter*. ~~etc...~~ This cell must have the capability of being transformable by any of the techniques known by the man skilled in the art. This can be, non-restrictively, transformation by electroporation, by conjugation, by transduction, or by infection~~, etc....~~. This cell is used as an expression host for individualising the vectorised fragments of DNA.

Please substitute the following paragraph on page 16, line 18:

Recombinant vector means an expression or cloning vector having integrated ~~an~~ with exogenous genetic information, for example genomic DNA or metagenomic DNA.

Please substitute the following paragraph on page 16, line 30:

Shuttle vector means a vector enabling the transfer and the maintenance ~~of a~~ of genetic information from one (or more) donor bacterial species or strain(s) to one or more host organism(s) or strain(s) or species.

Please substitute the following paragraph on page 17, line 16:

The environmental sample contains a multitude of organisms including eubacteria, archaebacteria, algae, fungi, yeasts, protozoans, viruses, phages, or parasites~~, etc~~. The microorganisms can be represented by extremophiles such as thermophiles, psychrophiles, acidophiles, halophiles, etc. The environmental sample can contain cultivatable or non-cultivatable, known or unknown microorganisms, as well as free nucleic acids and organic matter.

Please substitute the following paragraph on page 18, line 13:

The relative efficacy of these two approaches as well as their respective advantages and disadvantages have been the ~~objet~~ object of numerous scientific studies, and are known by the man

skilled in the art. Establishing the strategy for extracting the nucleic acids, *i.e.* the choice of one or other of these two approaches and the choice of the different methods, is based non-restrictively upon the characteristics of the environment being examined, upon the targeted microorganisms (all or some of the microorganisms of this environment) and their characteristics, upon the size of the nucleic acids, and upon the choice of the cloning vector and the cloning strategy chosen.

Please substitute the following paragraph on page 18, line 23 to page 19, line 2:

Any vector-host system known in the prior art can be used in this invention. The identification of a host cell forms **Step 1** of this invention. The host cell can be eukaryotic or prokaryotic. Preferably, the host cell used is a bacterial host. These can be, non-restrictively, *Escherichia coli*, *Bacillus subtilis*, *Streptomyces lividans*, *Pseudomonas*, or *Nocardia*, *Acinetobacter* ~~etc~~. Examples of eukaryotic host cells, without being restricted to these, are yeasts and fungi. The host cell (i) can originate from collections of public strains, from private laboratories or commercial companies ; (ii) must be selected or modified for its inability to metabolise the substrate(s) {Ai} and the target product {B} ; (iii) must be able to be cultivated in the standard conditions known by the man skilled in the art; (iv) must be capable of being transformed by any of the techniques known by the man skilled in the art ; (v) must finally stably maintain the transforming exogenous DNA despite possible systems such as recombination, restriction~~, etc ...~~.

Please substitute the following paragraph on page 19, line 10:

Numerous cloning or expression vectors have been described in the prior art. Non-restrictively, these are plasmids, cosmids such as those marketed by the companies Stratagene (SuperCos and pWE15) and Epicentre Technologies (pWeb cosmid cloning kit), fosmids as described by Kim *et al.* (1992, Nucl. Acids Res. 20 : 1083-1085), artificial chromosomes PAC as described by Ioannou *et al.*, (1994, Nat. Genet., 6 : 84-89), artificial chromosomes BAC as described by Shizuya *et al.* (1992, Proc. Natl. Acad. Sci., 89:8794-8797), artificial chromosomes YAC as described by Larin *et al.* (1991, Proc. Natl. Acad. Sci., 88 : 4123-4132), phagemids and vectors derived from phages such as those marketed by the company Stratagene (Lambda Dash II and Zap II), viral vectors~~, etc~~. Preferably, the vectors are of the cosmid, fosmid, BAC, YAC and P1-derivative

type because they enable the cloning of large fragments of DNA (between 30 kb and 200 kb and over for the BACs and the YACs). The vectors can either be integrative in that they integrate randomly or in a controlled way into the genome of the host cell, or preferably be replicative in that the vector is maintained in the host cell independently of the genome of this cell. By definition, the cloning vectors contain a certain number of elements necessary for maintaining the vector in the host cell (origin of functional replication), or else necessary for the selection and/or the detection of the vector in this cell (marker gene such as for example a resistance gene to an antibiotic under the functional promoter in the host cell and enabling a positive selection pressure). Due to the specificity of these constitutive elements, the vectors have a wider or narrower host spectrum.

Please substitute the following paragraphs on page 20, line 1:

The cloning process, *i.e.* the introduction of the sequences of nucleic acid, preferably purified metagenomic DNAs, into the appropriate vector, requires numerous steps of molecular manipulation of the DNAs (in a non-limitative way for the restrictions, dephosphorylations, ligations~~, etc.~~) which have been widely described, for example in Current Protocols in Molecular Biology, Eds. F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl, published by Greene Publishing Associates and Wiley Inter-Science. Two approaches for creating metagenomic libraries can be considered.

In a first preferred embodiment, the metagenomic library is formed directly in a **shuttle vector** specific of one or more hosts, preferably bacterial, for example as described in patents N° WO 01/40497A2 (Aventis Pharma, 1999) and W0 99/67374 (Biosearch Italia, 1999) for *Streptomyces*. In a second embodiment, the purified nucleic acids are cloned in a general vector, for example of the fosmid or BAC type, then the recombinant vectors are modified, individually or in a pool, advantageously by transposition as described in patent application ~~N°~~ No. PCT/EP 03/07765 (Libragen). In this process, the transposition makes it possible to introduce, either into the vector, or into the insert (disruption or activation), the genetic elements necessary for the transfer, the replication or the integration of the recombinant vector in the chosen host cell, preferably a bacterial host. This post-modification of the clones of the library can be implemented individually (metagenomic library structured in the format of 96 or 384 microplaques) or collectively (non-

structured metagenomic library). The transformation of the population of host cells identified in step 1) by a population of cloned DNAs forms **Step 2** of this invention. In the two embodiments, the metagenomic library can be structured in advance in that all of the clones of the library are individualised in a format capable of being automated (96, 384, 1536 microplaques) or preferably be preserved in the form of a mixture of recombinant clones. In this preferred preservation mode, the library can advantageously be amplified in that the host cells, after transformation or infection, are multiplied over a specific number of cycles, leading to every recombinant clone of the library being represented by n copies in the amplified library, and the amplified library being able to be subjected to numerous simultaneous screening or selection tests, without any loss of diversity.

Please substitute the following paragraph on page 22, line 29 to page 23, line 2:

- the mutated phenotype III (Ai- ; B-)* of transposed clones capable of using neither {Ai} nor {B} as sources of growth. Either (i) mutagenesis by transposition has reached an element common to the metabolic pathways of {Ai} and {B} such as for example a regulation element, a common transporter~~, etc.~~; or (ii) the metabolic pathway of {Ai} passes via {B} and the metabolic pathway of {B} enabling growth is disrupted.

Please substitute the following paragraph on page 23, line 11:

Step 8 : The genetic characterisation of the biocatalyst, *i.e.* characterisation of the gene or genes encoding the enzyme or enzymes involved in the conversion of {Ai} into {B}, is implemented by means of the transposed clones having the phenotype (Ai-B+). The genetic analysis of the nucleic sequences located on the disruption site or sites of the recombinant clones (Ai- ; B+) makes it possible to elucidate the genetic system(s) responsible for the conversion of {Ai} into {B}. The genetic analysis is implemented by any methods known by the man skilled in the art, including non-restrictively establishing sequences of nucleic acids, identifying coding and regulating sequences~~, etc~~.